Sifting Through The Usenet With The Netnews Filtering Service

(Or, Stanford 1, Savetz 0)

BY KEVIN SAVETZ

The Usenet is the world's largest distributed bulletin board system, shared by millions of people using hundreds of thousands of computers scattered along the Internet highway. At last count, there were over seven thousand Usenet "newsgroups," or discussion topics, covering every subject imaginable. Folks on the Usenet talk about everything you can think of, from square dancing to motorcycle maintenance, from the Swedish Chef to Ronald Reagan.

With so much going on on the Usenet, it had been impossible to keep abreast of it all. The best you could do was pick out the newsgroups that most interested you, read them as regularly as possible, and try not to let the ever-prolific Usenet keep you away from more important things, like regular meals. The biggest problem was you couldn't search all the newsgroups at once for the information that most interested you.

This changed in February with the introduction of the Stanford Netnews Filtering Service, the niftiest Internet hack in...well, in at least a month.

HOW IT WORKS

The Stanford Netnews Filtering Service is a tool for creating personalized delivery of Usenet news. You subscribe to the service by establishing "profiles" that describe your interests. The Filtering Service scans the vast quantity of goodies posted each day to the Usenet and emails you the articles that contain information you are interested in.

The best part is that this automated program searches all newsgroups (well, all those available at stanford.edu) for interesting articles. For instance, if you're interested in UFOs, and a conversation about them should pop up in the newsgroup "alt.fan.laurie-anderson" (an altogether unlikely place for such a conversation, admittedly), the filtering service will be sure you won't miss out, even if you don't normally read about Laurie Anderson.

Profiles are simple to set up because you just feed the filter plain English text, without boolean "ands," "ors," or "nots." For instance, you might search for "object-oriented programming" or "nba golden state warriors basketball." Based on the statistical distributions of the words in articles posted to the Usenet, each article is scored to evaluate how relevant it is to your profile. Articles that are deemed to be interesting to you are emailed to you.

For instance, if you're interested in UFOs, and a conversation about them should pop up in the newsgroup "alt.fan.laurie-anderson," the filtering service will be sure you won't miss out, even if you don't normally read about Laurie Anderson.

The service is rather robust; after you receive useful articles, you can feed the best ones back to the service to improve its search strategy. You can also adjust the frequency of delivery, the volume of articles and the length of your subscription to each profile.

A TRUE STORY

The following story is true. Only the names have been changed to protect the guilty. Once I discovered the Filtering Service, I entered a search for "Negativland" (an alternative music group out of Berkeley). I try to keep up with Negativland's exploits by reading "rec.music.alternative," but frankly, nothing else on that newsgroup interests me at all.

So I told the filtering service to look for articles about Negativland, set the search time limit for the rest of eternity, and happily waited. The following morning, two articles mentioning Negativland arrived in my email box. With a spring in my step, I unsubscribed from rec.music.alternative, content in the knowledge that I would remain informed about Negativland.

So far, so good. For my next trick, I started a news filter on the word "Savetz," my own last name. Throwing modesty to the wind, I decided that if anyone was talking about me, I ought to know about it.

It seems there's an obscure newsgroup called "air.unix". Its distribution is limited to a few computers on the second floor of Sweet Hall at Stanford University. It also seems that the Usenet Filter, since it lives at Stanford,

Copyright © 1994 by Kevin Savetz. All rights reserved. *Kevin Savetz is currently working on his first book,* Your Internet Consultant—The FAQS of Life Online, *due in June from Sams Publishing. His email address is Savetz@rahul.net.*

Efi's
DISCOUNT COMPUTERS

JUST A SAMPLE

Specializing in these quality products:
AST RESEARCH INC.
EVEREX
Canon
SAMSUNG
MITSUBISHI ELECTRONICS
TOSHIBA
intel
ZENITH
NEC
OMEGA
Makers of the Berserk Box
Acer

- NEW! Canon BubbleJet 200E, 2 year warranty $299
- AST Notebook 486SL25/ 33w/4MB, 200MB HD, Color, Windows, & Mouse 3795/3995
- Canon Laser LBP 430, 4PPM, 300dpi, 1MB, 2 year warranty 599
- Canon BubbleJet 600 color-unbelievable colors MUST SEE! 599
- AST Bravo 4/25S 730
- AST Bravo LC4/66, 200MB 1995
- NEC 5FGP Monitor 1189
- Texas Instruments Micro Laser Pro 600 w/6MB 1365
- Dual Speed AT CD-ROM w/ Sound Wave 32 High End Card and Four Titles 449

OKIDATA
IBM
COMPAQ
Advanced Logic Research, Inc.
ALPS AMERICA
Panasonic
EPSON
HEWLETT PACKARD
Hayes
CITIZEN

Fax for inquiries, price bids & purchase orders: (510) 283-1744
Free delivery and installation of systems within 30 miles. Free in-store installation of any product.
CALL US FOR YOUR PRICE ON ANY PRODUCT.
HOURS: MON-FRI. 9-5, SAT. 9-1
3441 GOLDEN GATE WAY, LAFAYETTE, CA 94549
(510) 283-1700
CALL BEFORE YOU BUY
We are actively seeking minority businesses supplying somputer products. Please Call!
Prices may change without notice. 90 day warranty on all products.

DISCOVER THE INTERNET
FOR ONLY
$17.50
A MONTH

Unlimited access with no hourly connect fees!
Why pay long distance phone charges?

CRL brings the Internet to you with a local call from any of the following areas:

•Alameda	•Emeryville	•Mill Valley	•Phoenix, AZ	•San Rafael
•Atlanta, GA	•Fairfax	•Mountain View	•Pinole	•Santa Rosa
•Anaheim	•Hercules	•New York, NY	•Redwood City	•Saratoga
•Belmont	•Ignacio	•Novato	•Richmond	•Sausalito
•Berkeley	•Kentfield	•Oakland	•Rohnert Park	•Sebastopol
•Boston, MA	•Los Altos	•Orinda	•San Carlos	•Sonoma
•Campbell	•Los Angeles	•Pacifica	•San Francisco	•St. Louis,MO
•Corte Madera	•Menlo Park	•Palo Alto	•San Jose	•Sunnyvale
•Crockett	•Millbrae	•Petaluma	•San Mateo	•Tiburon

CRL offers a full range of Internet services including: Telnet, FTP, E-Mail, Usenet News, Gopher, WAIS, IRC, MUDs, access to local CRL software archives, and 24 hour customer support. All dialups support SLIP/PPP, UUCP and Shell services at speeds up to 14400 baud. Contact CRL for your dialup or dedicated line service needs.

CRL
VOICE: 415-837-5300
FAX: 415-392-9000
E MAIL: support @ crl. com

can see "air.unix" and filter its articles, even though I, being a non-Stanford person, cannot. And, it seems that indeed, someone there was talking about me.

A woman at Stanford posted a message on "air.unix" asking how to send a fax from the Internet. Another soul, ever-helpful (although misguided about US copyright law), posted to the newsgroup the entire text of "Faxing on the Internet," an article I wrote for the February 7 issue of *MicroTimes*. He typed the entire fourteen hundred-word article in himself—I can only suppose to impress the woman who asked the question in the first place.

The Usenet Filter dutifully perused the article, found my name therein, and emailed it to me. To make a long story short (is it too late?), I told the "air.unix" folks a little about copyright law and gave them even more information about sending faxes from the Internet. I also told them about the Stanford Usenet Filtering Service, and explained that I had found their article because it contained my name. Were they grateful? Apparently not.

The next morning, I discovered that the filtering service was sending me lots of articles that I didn't care about at all. They didn't seem to be about Negativland and they didn't seem to mention me. This threw me for a loop, until I realized that every post from one of those wacky "air.unix" readers concluded with a phrase like "This line is just here so Kevin Savetz has to read this message." Ouch. Final score: Stanford, one. Me, zero.

There are two morals to this story. First, don't tell anyone if you use the filtering service to look for your own name. Second (and more importantly) please use the Internet only for good, not evil. ■

TO ACCESS THE SERVICE...

You can access the service from Mosaic: http://woodstock.stanford.edu:2000 .

The service also supports email access. To get the instructions on the email interface, send a message with the word "help" in the message body to netnews@db.stanford.edu .

Here is an example to give you some ideas of how the service works. Say you subscribe to the service with the profile "online information services." Then periodically you will receive email messages like this:

```
Subscription 1: online information services
 Article: misc.activism.progressive.11965
 From: hn0003@handsnet.org
 Subject: HandsNet WEEKLY DIGEST 1/15-21
 Score: 84
 First 15 lines:
 HANDSNET WEEKLY DIGEST January 15 - 21, 1994
 News from HandsNet's Information Forums
 HandsNet is a national, nonprofit network connecting
organizations working on social and economic justice
issues. Members use HandsNet to make new contacts,
work collaboratively and to find and publish informa-
tion, news ....
 Article: ca.politics.38420
 From: rlm@helen.surfcty.com (Robert L. McMillin)
 Subject: GOV-ACCESS #5:Cal.Emergency Svcs.online +
Net-fax + MINN Pub
Info Net
 Score: 82
 First 15 lines:
 Jan. 22, 1994
 CALIFORNIA OFFICE OF EMERGENCY SERVICES INFO AVAIL-
ABLE ONLINE
 <a recent exchange of messages>
 The state Emergency Digitial Information Service is
working fine Telnet to telnet oes1.oes.ca.gov 5501
```